The Nonlinear Dirac Equation in Bose-Einstein Condensates: Foundation and Symmetries

L. H. Haddad and L. D. Carr

Department of Physics, Colorado School of Mines, Golden, Colorado 80401, USA

Abstract

We show that Bose-Einstein condensates in a honeycomb optical lattice are described by a nonlinear Dirac equation in the long wavelength, mean field limit. Unlike nonlinear Dirac equations posited by particle theorists, which are designed to preserve the principle of relativity, i.e., Poincaré covariance, the nonlinear Dirac equation for Bose-Einstein condensates breaks this symmetry. We present a rigorous derivation of the nonlinear Dirac equation from first principles. We provide a thorough discussion of all symmetries broken and maintained.

1. Introduction

Recently the first truly two-dimensional (2D) solid state material, graphene, was created in the laboratory [1,2]. One of the most exciting aspects of this novel material is that long wavelength excitations are described by a Dirac equation for massless particles, with a "speed of light" equal to the Fermi velocity $v_F \simeq c/300$ [3]. Thus one can study relativistic phenomena at very low velocities in an experiment far more accessible than a particle accelerator. The only real requirement to obtain this equation is the simple hexagonal, or honeycomb lattice structure of the graphene [4]. One can therefore consider any solid state system constructed on a honeycomb lattice, including artificial systems, in order to study relativistic phenomena in novel materials accessible in tabletop experiments.

The most precise, cleanest, most controllable artificial solid state system is ultra-cold atoms in optical lattices. Such systems have no impurities and no disorder unless specifically added in by hand. They are very versatile: they can be constructed with an arbitrary lattice structure in one, two or three dimensions; they can contain bosons and/or fermions, atoms and/or diatomic molecules; and they can even have a pseudo-spin structure. Their tem-

perature and interaction sign, magnitude, and symmetry can be controlled externally. Moreover, 2D physics has recently been of great interest in this context, in the form of the Berzinskii-Kosterlitz-Thouless crossover [5], and 2D systems underpinned by lattices are immediately available in experiments. Instead of considering ultra-cold fermions, which could be used to produce an near-exact analog of graphene [6], we consider ultra-cold bosons. Bosonic statistics and interactions lead to a new feature in the massless Dirac equation known in graphene – a naturally occurring nonlinear term, giving rise to a *nonlinear Dirac equation.*

The study of nonlinear phenomena in ultra-cold atoms, especially in Bose-Einstein condensates (BECs) [7,8], has been enormously fruitful. The recent text edited by Kevrekidis, Frantzeskakis, and Carretero-González provides an excellent summary of this field [9]. The nonlinear mean field description given by the nonlinear Schrodinger equation (NLSE) has been very accurate in the majority of experiments on BECs. Vector and non-local generalizations of the NLSE have also proven useful [9]. In optical lattices, the mean field description remains accurate provided the lasers creating the standing wave which is the optical lattice are not too intense, and the dimensionality of the system is greater than one [10,11].

In this article, we present a completely new class of nonlinear phenomena in BECs, based on the nonlinear Dirac equation (NLDE). Nonlinear Dirac equations have a long

history in the literature, particularly in the context of particle and nuclear theory [12,13,14,15], but also in applied mathematics and nonlinear dynamics [16,17,18,19,20]. As nonlinearity is a ubiquitous aspect of Nature, it is natural to ask how nonlinearity might appear in a relativistic setting. However, this line of questioning has been strongly constrained by modeling, rather than first principles. That is, there is no standard first principle of quantum electrodynamics (QED) which is nonlinear. So, the approach has been to require symmetry constraints in nonlinear models. One of these constraints is the principle of relativity, i.e., Poincaré covariance. Poincaré symmetry includes rotations, translations, and Lorentz boosts.

In contrast, our NLDE is not a model: it is derived from first principles for a weakly interacting bosonic gas in the presence of a honeycomb optical lattice. We show that Poincaré symmetry is naturally broken by the nonlinearity inherent in this system. Given that this form of nonlinearity, which depends only on the local condensate density, is one of the most common throughout nature, it is important to recognize that the principle of relativity may be broken by small nonlinearities even at a fundamental level, for example of QED [21,22]. Thus, we suggest a new direction of investigation in particle physics of possible nonlinearities as well as providing a natural context in artificial solid state systems for the NLDE.

This article is outlined as follows. In Sec. 2 we provide a rigorous derivation of the NLDE from first principles. In Sec. 3 we discuss both discrete and continuous symmetries common to relativistic systems, providing a clear physical interpretation in the present context. Finally, in Sec. 4 we conclude.

2. The nonlinear Dirac equation

2.1. Two-Component Spinor Form of the NLDE

The second quantized Hamiltonian for a weakly interacting bosonic gas in two spatial dimensions is

$$\hat{H} = \int d^2r \, \hat{\psi}^\dagger H_0 \hat{\psi} + \frac{g}{2} \int d^2r \, \hat{\psi}^\dagger \hat{\psi}^\dagger \hat{\psi} \hat{\psi} \,, \tag{1}$$

$$H_0 \equiv -\frac{\hbar^2}{2m} \nabla^2 + V(\vec{r}) \,. \tag{2}$$

The bosonic field operators $\hat{\psi} = \hat{\psi}(\vec{r}, t)$ obey bosonic commutation relations in the Heisenberg picture. In Eq. (1), $g \equiv 4\pi\hbar^2 a_s/m$ is the coupling strength for binary contact interactions with a_s the s-wave scattering length and m the atomic mass. The external potential $V(\vec{r})$ is a honeycomb lattice formed by standing waves of three sets of counter-propagating laser beams [6]. The atoms experience this potential via the AC Stark effect. We assume that the third spatial dimension is frozen out by a tightly confining potential which is locally harmonic, as in Ref. [5].

The honeycomb lattice has two sites in the lattice unit cell. We refer to the resulting two degenerate sublattices as

(a) Hexagonal lattice structure

(b) Reciprocal lattice
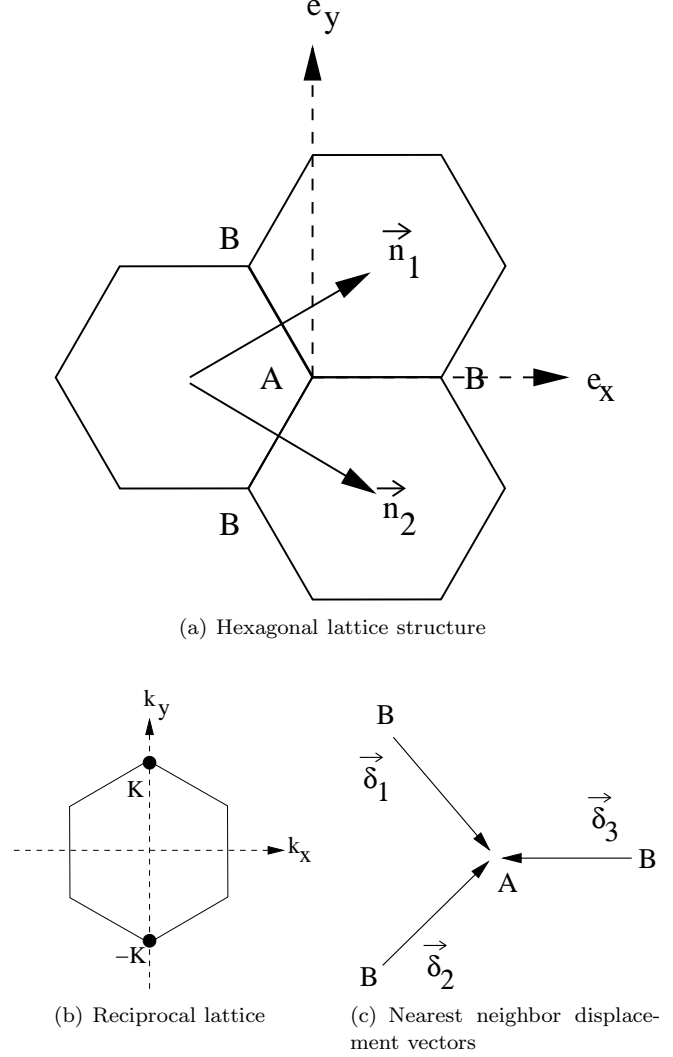
(c) Nearest neighbor displacement vectors

Fig. 1. Characterization of a honeycomb lattice.

A and B. Expanding in terms of Bloch states in the lowest band belonging to A or B sites of the honeycomb lattice, as shown in Fig. 1, we can break up the bosonic field operator into a sum over the two sublattices:

$$\hat{\psi} = \hat{\psi}_A + \hat{\psi}_B \,, \tag{3}$$

$$\hat{\psi}_A \equiv \sum_A \hat{a} \, e^{i\vec{k}\cdot(\vec{r}-\vec{r}_A)} u(\vec{r} - \vec{r}_A) \,, \tag{4}$$

$$\hat{\psi}_B \equiv \sum_B \hat{b} \, e^{i\vec{k}\cdot(\vec{r}-\vec{r}_B)} u(\vec{r} - \vec{r}_B) \,, \tag{5}$$

where $\hat{a}$ and $\hat{b}$ are the time-dependent destruction operators at A and B sites and $\vec{r}_A$ and $\vec{r}_B$ are the positions of A and B sites, respectively. The spatial dependence is then encapsulated outside the operator in the exponential and the functions u. The summation indices indicate sums over A or B sites.

Inserting Eq. (3) into Eq. (1), the Hamiltonian can be rewritten

$$\hat{H} = \int d^2r \left[(\hat{\psi}_A^\dagger + \hat{\psi}_B^\dagger) H_0 (\hat{\psi}_A + \hat{\psi}_B) \right.$$

$$+\frac{g}{2}(\hat{\psi}_A^\dagger + \hat{\psi}_B^\dagger)(\hat{\psi}_A^\dagger + \hat{\psi}_B^\dagger)(\hat{\psi}_A + \hat{\psi}_B)(\hat{\psi}_A + \hat{\psi}_B)\Big] \ . \quad (6)$$

In the integral over H_0, imposing the restriction of nearest-neighbor interactions in the tight-binding, lowest band approximation eliminates all A-A and B-B transitions except for on-site kinetic and potential terms; the latter can be neglected as an overall self-energy. Then only integrals involving neighboring A-B sites remain in the sum. Similarly, in the interaction term only on-site terms are non-negligible, i.e., overlap of functions u belonging to the same site. Thus in the tight-binding, lowest band approximation, Eqs. (4)-(5) are substituted into Eq. (6) to yield:

$$\hat{H} = \int d^2r \sum_{<A,B>} \Big[\hat{a}^\dagger\, e^{-i\vec{k}\cdot\vec{\chi}_A} u(\vec{\chi}_A) H_0 \hat{a}\, e^{i\vec{k}\cdot\vec{\chi}_A} u(\vec{\chi}_A)$$
$$+\hat{a}^\dagger\, e^{-i\vec{k}\cdot\vec{\chi}_A} u(\vec{\chi}_A) H_0 \hat{b}\, e^{i\vec{k}\cdot\vec{\chi}_B} u(\vec{\chi}_B)$$
$$+\hat{b}^\dagger\, e^{-i\vec{k}\cdot\vec{\chi}_B} u(\vec{\chi}_B) H_0 \hat{a}\, e^{i\vec{k}\cdot\vec{\chi}_A} u(\vec{\chi}_A)$$
$$+\hat{b}^\dagger\, e^{-i\vec{k}\cdot\vec{\chi}_B} u(\vec{\chi}_B) H_0 \hat{b}\, e^{i\vec{k}\cdot\vec{\chi}_B} u(\vec{\chi}_B)\Big]$$
$$+\frac{g}{2}\int d^2r \sum_A \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a}\,[u(\vec{\chi}_A)]^4$$
$$+\frac{g}{2}\int d^2r \sum_B \hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{b}\,[u(\vec{\chi}_B)]^4 \ , \quad (7)$$

$$\vec{\chi}_A \equiv \vec{r} - \vec{r}_A\,, \quad \vec{\chi}_B \equiv \vec{r} - \vec{r}_B\,, \quad \vec{\chi}_{AB} = \vec{r}_A - \vec{r}_B\,. \quad (8)$$

Here the bracketed A and B summation index signifies a sum over nearest-neighbor A and B sites. Isolating the integrals by pulling out all sums and terms not dependent on $\vec{r}$, Eq. (7) becomes

$$\hat{H} = \sum_{<A,B>} \Big[\hat{a}^\dagger\hat{b}\, e^{i\vec{k}\cdot\vec{\chi}_{AB}} \int d^2r\, e^{-i\vec{k}\cdot\vec{r}} u(\vec{\chi}_A) H_0 e^{i\vec{k}\cdot\vec{r}} u(\vec{\chi}_B)$$
$$+ \hat{b}^\dagger\hat{a}\, e^{-i\vec{k}\cdot\vec{\chi}_{AB}} \int d^2r\, e^{-i\vec{k}\cdot\vec{r}} u(\vec{\chi}_B) H_0 e^{i\vec{k}\cdot\vec{r}} u(\vec{\chi}_A) \Big]$$
$$+\sum_A \hat{a}^\dagger\hat{a} \int d^2r\, e^{-i\vec{k}\cdot\vec{r}} u(\vec{\chi}_A) H_0 e^{i\vec{k}\cdot\vec{r}} u(\vec{\chi}_A)$$
$$+\sum_B \hat{b}^\dagger\hat{b} \int d^2r\, e^{-i\vec{k}\cdot\vec{r}} u(\vec{\chi}_B) H_0 e^{i\vec{k}\cdot\vec{r}} u(\vec{\chi}_B)$$
$$+\frac{g}{2}\sum_A \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} \int d^2r\,[u(\vec{\chi}_A)]^4$$
$$+\frac{g}{2}\sum_B \hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{b} \int d^2r\,[u(\vec{\chi}_B)]^4 \ , \quad (9)$$

Finally, we redefine the spatial integrals in Eq. (9) as hopping energy t_h and interaction energy U, respectively, as is standard for the Hubbard Hamiltonian [23,24] (note that we reserve t for time). The terms in Eq. (9) proportional to $\hat{a}^\dagger\hat{a}$ and $\hat{b}^\dagger\hat{b}$ just count the total number of atoms in the system, and can been neglected as an overall constant. This leads to the Hamiltonian

$$\hat{H} = -t_h \sum_{<A,B>} \Big[\hat{a}^\dagger\hat{b}\, e^{i\vec{k}\cdot(\vec{r}_A - \vec{r}_B)} + \hat{b}^\dagger\hat{a}\, e^{-i\vec{k}\cdot(\vec{r}_A - \vec{r}_B)} \Big]$$

$$+\frac{U}{2}\sum_A \hat{a}^\dagger\hat{a}^\dagger\hat{a}\hat{a} + \frac{U}{2}\sum_B \hat{b}^\dagger\hat{b}^\dagger\hat{b}\hat{b}\,. \quad (10)$$

Equation (10) is the Hubbard Hamiltonian divided into two degenerate sublattices A and B, appropriate to the honeycomb optical lattice.

In order to work towards the nonlinear Dirac equation, we calculate the time evolution of $\hat{a}$ and $\hat{b}$ according to the standard Heisenberg picture prescription. This is similar to the approach taken by Pitaevskii in his landmark paper which first obtained the NLSE, or Gross-Pitaevskii equation [25]. The Heisenberg equation of motion is

$$i\hbar\,\partial_t\hat{a}_k = [\hat{a}_k, \hat{H}]\,. \quad (11)$$

The operator $\hat{a}_k$, which destroys a boson at site k on sublattice A, satisfies the bosonic commutation relation

$$[\hat{a}_k, \hat{a}_{k'}^\dagger] = \delta_{kk'}\,. \quad (12)$$

Then the commutator with the on-site interaction terms reduces to

$$[\hat{a}_k, \hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k] = \hat{a}_k\hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k - \hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k\hat{a}_k\,. \quad (13)$$

Taking the first product on the right and commuting the furthermost left $\hat{a}$ through according to Eq. (12), one finds:

$$\hat{a}_k\hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k = 2\hat{a}_k^\dagger\hat{a}_k\hat{a}_k + \hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k\hat{a}_k\,. \quad (14)$$

Substituting Eq. (14) into Eq. (13), one obtains

$$[\hat{a}_k, \hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k\hat{a}_k] = 2\hat{a}_k^\dagger\hat{a}_k\hat{a}_k\,. \quad (15)$$

Substituting Eq. (15) into Eq. (11) and the Hubbard Hamiltonian Eq. (10), one finds

$$i\hbar\,\partial_t\hat{a}_k = -t_h \Big[\hat{b}_k\, e^{i\vec{k}\cdot(\vec{r}_{A_k}-\vec{r}_{B_k})} + \hat{b}_{k-n_1}\, e^{i\vec{k}\cdot(\vec{r}_{A_k}-\vec{r}_{B_{k-n_1}})}$$
$$+\hat{b}_{k-n_2}\, e^{i\vec{k}\cdot(\vec{r}_{A_k}-\vec{r}_{B_{k-n_2}})} \Big] + U\hat{a}_k^\dagger\hat{a}_k\hat{a}_k \quad (16)$$

where the first three terms on the right hand side represent transitions from the three B-sites nearest the kth site of the A sublattice and $\vec{n}_1$ and $\vec{n}_2$ are primitive cell translation vectors for the reciprocal lattice, as shown in Fig. 1.

In a similar fashion as Eqs. (11)-(16), we arrive at an expression of the same form for the B sublattice:

$$i\hbar\,\partial_t\hat{b}_k = -t_h \Big[\hat{a}_k\, e^{-i\vec{k}\cdot(\vec{r}_{A_k}-\vec{r}_{B_k})} + \hat{a}_{k+n_1}\, e^{-i\vec{k}\cdot(\vec{r}_{A_{k+n_1}}-\vec{r}_{B_k})}$$
$$+\hat{a}_{k+n_2}\, e^{-i\vec{k}\cdot(\vec{r}_{A_{k+n_2}}-\vec{r}_{B_k})} \Big] + U\hat{b}_k^\dagger\hat{b}_k\hat{b}_k\,. \quad (17)$$

Continuing to follow Pitaevskii's method, we next take the expectation value of Eqs. (16) and (17) with respect to on-site coherent states. A tensor product over sites of such coherent states is also assumed [10]. A more formal, careful treatment of finite number states, rather than coherent states, has been worked out in the literature (see [26] and references therein). Either way, we obtain coupled equations of motion for discrete, on-site, complex-valued amplitudes. For simplicity of notation we take

$$a_k \equiv \langle\hat{a}_k\rangle\,, \quad b_k \equiv \langle\hat{b}_k\rangle\,. \quad (18)$$

3

Inserting the nearest-neighbor vectors $\vec{\delta}_1$, $\vec{\delta}_2$, and $\vec{\delta}_3$ in the exponentials in Eqs. (16) and (17), as shown in Fig. 1, we obtain

$$i\hbar \dot{a}_k = -t_h(b_k e^{i\vec{k}\cdot\vec{\delta}_3} + b_{k-n_1} e^{i\vec{k}\cdot\vec{\delta}_1} + b_{k-n_2} e^{i\vec{k}\cdot\vec{\delta}_2})$$
$$+U a_k^* a_k a_k , \tag{19}$$
$$i\hbar \dot{b}_k = -t_h(a_k e^{-i\vec{k}\cdot\vec{\delta}_3} + a_{k+n_1} e^{-i\vec{k}\cdot\vec{\delta}_1} + a_{k+n_2} e^{-i\vec{k}\cdot\vec{\delta}_2})$$
$$+U b_k^* b_k b_k , \tag{20}$$

where n_1 and n_2 in the indices label the lattice sites in the two directions of the primitive-cell translation vectors $\vec{n}_1$ and $\vec{n}_2$.

The NLDE is derived around the linear band crossings between the A and B sublattices at the Brillouin zone corners [27], called *Dirac cones* in the graphene literature [3]. To this end, we insert particular values for the nearest-neighbor displacement vectors $\vec{\delta}$ and evaluate $\vec{k}$ at the Brillouin zone corner, defined by $\vec{k} = \vec{K} = (0, 4\pi/3)$, $\vec{\delta}_1 = (\frac{1}{2\sqrt{3}}, -\frac{1}{2})$, $\vec{\delta}_2 = (\frac{1}{2\sqrt{3}}, \frac{1}{2})$, $\vec{\delta}_3 = (-\frac{1}{\sqrt{3}}, 0)$, all in the $x-y$ plane. Then Eq. (19) becomes

$$i\hbar \dot{a}_k = -t_h(b_k e^0 + b_{k-n_1} e^{-i2\pi/3} + b_{k-n_2} e^{i2\pi/3})$$
$$+U a_k^* a_k a_k . \tag{21}$$

Reducing the exponentials,

$$i\hbar \dot{a}_k = -t_h[b_k + b_{k-n_1}(-1/2 - i\sqrt{3}/2)$$
$$+b_{k-n_2}(-1/2 + i\sqrt{3}/2)] + U a_k^* a_k a_k . \tag{22}$$

In anticipation of taking the long wavelength, continuum limit, as is necessary to obtain the NLDE, we group terms appropriately in Eq. (22) in order to construct discrete versions of derivatives:

$$i\hbar \dot{a}_k = -t_h[b_k + (b_k - b_{k-n_1})(1/2 + i\sqrt{3}/2)$$
$$-b_k(1/2 + i\sqrt{3}/2) + (b_k - b_{k-n_2})(1/2 - i\sqrt{3}/2)$$
$$-b_k(1/2 - i\sqrt{3}/2)] + U a_k^* a_k a_k , \tag{23}$$

which reduces to

$$i\hbar \dot{a}_k = -t_h[(b_k - b_{k-n_1})(1/2 + i\sqrt{3}/2)$$
$$+(b_k - b_{k-n_2})(1/2 - i\sqrt{3}/2)] + U a_k^* a_k a_k . \tag{24}$$

Taking the continuum limit and replacing the discrete quantities $a_k = a_k(t)$ and $b_k = b_k(t)$ by the continuous functions $\psi_A = \psi_A(\vec{r}, t)$ and $\psi_B = \psi_B(\vec{r}, t)$, we arrive at

$$i\hbar \dot{\psi}_A = -t_h \left[\partial_{n_1} \psi_B(1/2 + i\sqrt{3}/2) + \partial_{n_2}(1/2 - i\sqrt{3}/2) \right]$$
$$+U \psi_A^* \psi_A \psi_A . \tag{25}$$

where the partial derivatives are in the directions of the unit-cell vectors $\vec{n}_1$ and $\vec{n}_2$. With a little trigonometry we find that the unit-cell vectors are

$$\vec{n}_1 = \cos(\pi/6)\hat{e}_x - \sin(\pi/6)\hat{e}_y = \sqrt{3}/2\hat{e}_x - 1/2\hat{e}_y , \tag{26}$$
$$\vec{n}_2 = \cos(\pi/6)\hat{e}_x + \sin(\pi/6)\hat{e}_y = \sqrt{3}/2\hat{e}_x + 1/2\hat{e}_y . \tag{27}$$

Up to now the "hat" symbol (circumflex accent) has been reserved for operators alone. However, in Eqs. (26)-(27) we use this symbol to indicate a unit vector in the x and y directions. Derivatives with respect to the unit-cell vectors take the form

$$\partial_{n_1} = \vec{n_1} \cdot \vec{\nabla} = (\sqrt{3}/2)\partial_x - (1/2)\partial_y , \tag{28}$$
$$\partial_{n_2} = \vec{n_2} \cdot \vec{\nabla} = (\sqrt{3}/2)\partial_x + (1/2)\partial_y . \tag{29}$$

Substituting Eqs. (28)-(29) into Eq. (25),

$$i\hbar \dot{\psi}_A = -t_h \Big\{ [(\sqrt{3}/2)\partial_x - (1/2)\partial_y]\psi_B[(1/2) + i(\sqrt{3}/2)]$$
$$+[(\sqrt{3}/2)\partial_x + (1/2)\partial_y]\psi_B[(1/2) - i(\sqrt{3}/2)] \Big\}$$
$$+U \psi_A^* \psi_A \psi_A . \tag{30}$$

Further simplification of Eq. (30) leads to

$$i\hbar \dot{\psi}_A = -\frac{t_h\sqrt{3}}{2}(\partial_x \psi_B - i\partial_y \psi_B) + U \psi_A^* \psi_A \psi_A . \tag{31}$$

Similarly, for the continuum limit of $b_k = b_k(t) \rightarrow \psi_B = \psi_B(\vec{r}, t)$,

$$i\hbar \dot{\psi}_B = -\frac{t_h\sqrt{3}}{2}(-\partial_x \psi_A - i\partial_y \psi_A) + U \psi_B^* \psi_B \psi_B . \tag{32}$$

Equations (31)-(32) are in fact massless Dirac equations with an added nonlinear term. To put this in a more familiar form, we use Pauli matrix notation. To this end, one must rename the coordinate axes so that $x \rightarrow y$, and, in order to preserve the handedness of the coordinate system, $y \rightarrow -x$. We also reinsert the lattice constant a; note that a is unrelated to the s-wave scattering length a_s briefly mentioned in the definition of g following Eq. (1). Thus $\partial_x \rightarrow \partial_y$ and $\partial_y \rightarrow -\partial_x$. Then Eqs. (31)-(32) become

$$i\hbar \dot{\psi}_A = -\frac{t_h a\sqrt{3}}{2}(\partial_y \psi_B + i\partial_x \psi_B) + U \psi_A^* \psi_A \psi_A , \tag{33}$$

$$i\hbar \dot{\psi}_B = -\frac{t_h a\sqrt{3}}{2}(-\partial_y \psi_A + i\partial_x \psi_A) + U \psi_B^* \psi_B \psi_B , \tag{34}$$

or in matrix form,

$$i\hbar \begin{bmatrix} \dot{\psi}_A \\ \dot{\psi}_B \end{bmatrix} = \frac{-it_h a\sqrt{3}}{2} \begin{bmatrix} 0 & \partial_x - i\partial_y \\ \partial_x + i\partial_y & 0 \end{bmatrix} \begin{bmatrix} \psi_A \\ \psi_B \end{bmatrix}$$
$$+U \begin{bmatrix} \psi_A^* \psi_A \psi_A \\ \psi_B^* \psi_B \psi_B \end{bmatrix} . \tag{35}$$

We can write Eq. (35) more compactly in terms of Pauli matrices $(\sigma_x, \sigma_y) = \vec{\sigma}$,

$$i\hbar \begin{bmatrix} \dot{\psi}_A \\ \dot{\psi}_B \end{bmatrix} = \frac{-it_h a\sqrt{3}}{2} \vec{\sigma} \cdot \vec{\nabla} \begin{bmatrix} \psi_A \\ \psi_B \end{bmatrix} + U \begin{bmatrix} \psi_A^* \psi_A \psi_A \\ \psi_B^* \psi_B \psi_B \end{bmatrix} . \tag{36}$$

Equation (36) is the NLDE in (2+1) dimensions.

4

However, we can make one further step by expressing Eq. (36) in a more covariant-looking form as follows:

$$(i\sigma_0\partial_t + ic_s\vec{\sigma}\cdot\vec{\nabla})\begin{bmatrix}\psi_A \\ \psi_B\end{bmatrix} - U\begin{bmatrix}\psi_A^*\psi_A\psi_A \\ \psi_B^*\psi_B\psi_B\end{bmatrix} = 0\,, \qquad (37)$$

where $\vec{\sigma}$ and $\vec{\nabla}$ are restricted to the $x-y$ plane. In Eq. (37)

$$c_s \equiv \frac{t_h a\sqrt{3}}{2\hbar} \qquad (38)$$

is an effective speed of sound of the condensate in the lattice[28,29] (in graphene it would be replaced with the Fermi velocity [3]). This velocity is an effective speed of light for excitations of the NLDE in our QED$_{2+1}$ theory. Experimental values of c_s in BECs are on the order of cm/s, ten orders of magnitude slower than the speed of light in a vacuum. Note also that in Eq. (37) U has now absorbed a factor of $1/\hbar$.

Finally, a few additional definitions lead to a nicely compact form for the NLDE. Let

$$A \equiv \begin{bmatrix}1 & 0 \\ 0 & 0\end{bmatrix}, B \equiv \begin{bmatrix}0 & 0 \\ 0 & 1\end{bmatrix}, \psi \equiv \begin{bmatrix}\psi_A \\ \psi_B\end{bmatrix}, \bar{\psi} \equiv \begin{bmatrix}\psi_A^*, \psi_B^*\end{bmatrix}. \quad (39)$$

With the choice of metric which raises and lowers spacetime indices restricted to the (2+1) dimensions,

$$g^{\mu\nu} = \begin{bmatrix}1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1\end{bmatrix}\,, \qquad (40)$$

Eq. (37) becomes

$$(i\sigma^\mu\partial_\mu - U\bar{\psi}A\psi A - U\bar{\psi}B\psi B)\psi = 0\,, \qquad (41)$$

where the standard Einstein summation rule is in effect, $\mu \in \{0,1,2\}$ in keeping with (2+1) dimensions, and the units are chosen such that $c_s = 1$.

2.2. Maximally Compact Form of the NLDE

In Sec. 2.1 we developed ψ, a two-dimensional complex object which brings to mind one member of a pair of Weyl-spinors in the (1/2,1/2) chiral representation of the Dirac algebra used to describe massless neutrinos in the standard QED$_{3+1}$ theory. Such a treatment is appropriate for any neutral Dirac fermion viewed in the extreme relativistic frame. In order to make the connection clear we must find the second member of the pair of Weyl-spinors and verify that the mapping is true. A thorough treatment of the mapping of QED$_{3+1}$ into the QED$_{2+1}$ theory of graphene is contained in [30] and references therein. We will continue to restrict ourselves to (2+1) dimensions in the following.

To this end, we seek to put the NLDE into a form consistent with the standard compact four-component spinor notation for the linear Dirac equation

$$i\gamma^\mu\partial_\mu\Psi = 0\,, \qquad (42)$$

where the γ^μ are the 4×4 Dirac matrices [31]. We point out that this notation is not only standard but more appropriate if the quasi-particles, i.e., the long-wavelength excitations, develop a non-zero effective mass, as can be caused by lattice distortion [6].

We obtained the NLDE by evaluating the exponentials at the Brillouin Zone corner $\vec{K}_+ = (0, 4\pi/3)$. There is another inequivalent corner, as shown in Fig. 1, $\vec{K}_- \equiv (0, -4\pi/3)$, near which perturbations in momentum, i.e., long-wavelength quasiparticle excitations, are governed by a similar first-order wave equation. By considering equations of motions derived around both Brillouin corners, we can obtain the four-vector notation. The coupled equations evaluated at $\vec{K}_-$ are

$$i\hbar\dot{\psi}_A = \frac{-t_h a\sqrt{3}}{2}(\partial_x + i\partial_y)\psi_B + U\psi_A^*\psi_A\psi_A\,, \qquad (43)$$

$$i\hbar\dot{\psi}_B = \frac{-t_h a\sqrt{3}}{2}(-\partial_x + i\partial_y)\psi_A + U\psi_B^*\psi_B\psi_B\,. \qquad (44)$$

Following the same steps as before we obtain

$$i\hbar\begin{bmatrix}\dot{\psi}_B \\ \dot{\psi}_A\end{bmatrix} = \frac{it_h a\sqrt{3}}{2}\vec{\sigma}\cdot\vec{\nabla}\begin{bmatrix}\psi_B \\ \psi_A\end{bmatrix} + U\begin{bmatrix}\psi_B^*\psi_B\psi_B \\ \psi_A^*\psi_A\psi_A\end{bmatrix}. \quad (45)$$

We combine Eqs. (45) and (36) into one equation involving a single 4-component object and attach $\pm$ subscripts to the wave functions Ψ to specify the corner of the Brillouin Zone. The resulting expression is

$$i\partial_t\begin{bmatrix}\Psi_+ \\ \Psi_-\end{bmatrix} + ic_s\begin{bmatrix}\vec{\sigma}\cdot\vec{\nabla} & 0 \\ 0 & -\vec{\sigma}\cdot\vec{\nabla}\end{bmatrix}\begin{bmatrix}\Psi_+ \\ \Psi_-\end{bmatrix}$$
$$-U\begin{bmatrix}N_+ \\ N_-\end{bmatrix} = 0\,. \qquad (46)$$

where the nonlinear terms are grouped into the 2-vectors N_+ and N_- defined by

$$N_+ = \begin{bmatrix}(\psi_A^*\psi_A\psi_A)_+ \\ (\psi_B^*\psi_B\psi_B)_+\end{bmatrix}, \quad N_- = \begin{bmatrix}(\psi_B^*\psi_B\psi_B)_- \\ (\psi_A^*\psi_A\psi_A)_-\end{bmatrix} \qquad (47)$$

and the four-spinors are given by

$$\Psi \equiv \begin{bmatrix}\Psi_+ \\ \Psi_-\end{bmatrix} \equiv \begin{bmatrix}\psi_{A+} \\ \psi_{B+} \\ \psi_{B-} \\ \psi_{A-}\end{bmatrix}, \qquad (48)$$

$$\Psi^\dagger \equiv \begin{bmatrix}\Psi_+^* & \Psi_-^*\end{bmatrix} \equiv \begin{bmatrix}\psi_{A+}^* & \psi_{B+}^* & \psi_{B-}^* & \psi_{A-}^*\end{bmatrix}. \qquad (49)$$

5

Here again the $\pm$ subscripts refer to the specific corner of the Brillouin zone.

We can reduce Eq. (46) to a more compact form by introducing the following 4×4 matrices:

$$\Sigma_0 = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}, \Sigma_1 = \begin{bmatrix} \sigma_x & \mathbf{0} \\ \mathbf{0} & -\sigma_x \end{bmatrix}, \Sigma_2 = \begin{bmatrix} \sigma_y & \mathbf{0} \\ \mathbf{0} & -\sigma_y \end{bmatrix}, \quad (50)$$

$$A_+ = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, A_- = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} \end{bmatrix}, \quad (51)$$

$$B_+ = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, B_- = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}. \quad (52)$$

The boldface notation $\mathbf{A}$ and $\mathbf{B}$ denote the 2×2 matrices defined in Eq. (39). Also, the boldface entries $\mathbf{1}$ and $\mathbf{0}$ in the matrices in Eq. (47) refer to the 2×2 unit matrix and zero matrix, respectively [15]. Then Eq. (46) becomes

$$(i\Sigma^\mu \partial_\mu - U\Psi^\dagger A_+ \Psi A_+ - U\Psi^\dagger B_+ \Psi B_+$$
$$- U\Psi^\dagger A_- \Psi A_- - U\Psi^\dagger B_- \Psi B_-)\Psi = 0. \quad (53)$$

We substitute into Eq. (53) the Dirac matrices in the Chiral representation:

$$\gamma^0 \equiv \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{bmatrix}, \gamma^1 \equiv \begin{bmatrix} \mathbf{0} & -\sigma_x \\ \sigma_x & \mathbf{0} \end{bmatrix}, \gamma^2 = \begin{bmatrix} \mathbf{0} & -\sigma_y \\ \sigma_y & \mathbf{0} \end{bmatrix} \quad (54)$$

and multiply on the left by γ^0. Note that $\mu \in \{0, 1, 2\}$ in keeping with $(2+1)$ dimensions. Finally, for the nonlinear term we introduce

$$N \equiv -U\Psi^\dagger \sum_{Q \in A_+, A_-, B_+, B_-} Q\Psi Q. \quad (55)$$

Implementing Dirac matrix notation as described, the NLDE of Eq. (53) becomes

$$(i\gamma^\mu \partial_\mu + \gamma^0 N)\Psi = 0. \quad (56)$$

Equation (56) is the most compact form of the massless NLDE in 4-component spinor notation.

3. Symmetries and Constraints

The NLDE as expressed by Eq. (37) or Eq. (41) looks like the Dirac equation for a two-spinor, as stated in the graphene problem [3], with the addition of two nonlinear on-site interaction terms, one for each A and B sublattice; similarly, the more compact form developed in Sec. 2.2 also appears to be a Dirac equation for a four-component spinor, with an additional term. However, one should not be too hasty in assigning characteristics based on appearances. We therefore make a careful and thorough exploration of the symmetries and other important mathematical properties of the NLDE.

In what follows we follow a similar route as in Ref. [15]. First we check the linear version with the delta interactions turned off to ensure that Eq. (41) is indeed the massless Dirac equation in $(2+1)$ dimensions with all the necessary symmetries. For each symmetry we then check the nonlinear interaction terms to determine whether they preserve or break the symmetry.

3.1. Locality

We do not necessarily require the evolution of the wave-function as described by an NLDE to be governed by a local theory. A local theory is one in which the terms in the linear equations of motion involve only factors of the wave-function and its derivatives evaluated at the same space-time coordinate. Nonlocality arises in low energy limits of some quantum field theories. However, the nonlinearity in our NLDE is manifestly local. Thus our NLDE is closer to the standard Dirac equation on the classical level (no quantum effects), modified by the on-site interaction term. In Sec. 4 we discuss the possibility of non-local nonlinearities, including for graphene.

3.2. Poincaré Symmetry

A Poincaré transformation takes the spatiotemporal point defined by the the 4-vector r_ν into the point r'_μ according to

$$r'_\mu = \Lambda^\nu_\mu r_\nu + d_\mu \quad (57)$$

where Λ^ν_μ is the coordinate matrix representation of the Lorentz group and d_μ is a space-time translation. The wave function Ψ transforms as

$$\Psi'(r') = M(\Lambda)\Psi(r)$$

where the matrices $M(\Lambda)$ form a representation of the subgroup of the Lorentz group consisting of spatial rotations and boosts. Boosts can be thought of as rotations in imaginary time by imaginary angles mixing space and time coordinates. We restrict these transformations to $(2+1)$ dimensions.

The proof of Lorentz covariance of the standard massless Dirac equation, Eq. (42), is arrived at with the aid of the transformations for the wave function and partial derivatives:

$$\Psi(r) = M^{-1}(\Lambda)\Psi'(\mathbf{r}'), \quad (58)$$
$$\partial_{x_i} = \Lambda^j_i \partial_{x'_j}. \quad (59)$$

This yields the conditions for the Dirac matrices:

$$\gamma_j = \Lambda_{ji} M \gamma_i M^{-1}. \quad (60)$$

The standard form of the Dirac matrices obtained this way can be found in the literature [31] and are identical to the results of our theory.

6

Thus imposing Lorentz covariance on the NLDE as expressed in Eq. (56) requires Ψ to transform under the irreducible representation of a subgroup of SL(2,C), the 2×2 complex matrices of unit determinant. The four-dimensional representation of SL(2,C), $D^{(1/2,1/2)}$, is formed by taking the direct product of the two-dimensional representations $D^{(1/2,0)}$ and $D^{(0,1/2)}$:

$$D^{(1/2,1/2)} = D^{(1/2,0)} \otimes D^{(0,1/2)} . \tag{61}$$

This subgroup of SL(2,C) is isomorphic to a subgroup of the Lorentz group, the one obtained by restricting Lorentz transformations to the plane of the honeycomb lattice. So, the upper two components of Ψ transform as a spinor under $D^{(1/2,0)}$, the lower two as a spinor under $D^{(0,1/2)}$, and Ψ itself as a four-component spinor or *bispinor* [32].

The next task is to examine the behavior of Ψ, as defined in Eq. (49) and governed by Eq. (56), under rotations in the $x - y$ plane. In order to obtain Poincaré covariance we must show that it is the same as that of a four-component spinor in the standard Dirac$_{3+1}$ theory restricted to the 2D plane. The honeycomb lattice is invariant under rotations by $\pm 2\pi/3$ but the four components of Ψ are also defined by the particular corner of the Brillouin Zone. Since we're considering a discrete lattice it is natural to discuss discrete rotations which realign lattice points and in the continuum limit map the continuous rotations of QED$_{2+1}$ onto our theory. Rotations by $\pm\pi/3$ exchange A and B sites, and take the theory to that of the opposite $\vec{K}$ point: $\vec{K}_+$ does not go to $\vec{K}_-$ but the result after calculating the relative phase exponentials gives back the same theory. To see this we chose a different primitive unit cell, the one obtained by a rotation of $2\pi/3$ about the $\vec{n}_1, \vec{n}_2$ origin in Fig 1. This is because the direction of $\vec{K}'_+$ (defined as $\vec{K}_+$ rotated by $\pi/3$) differs from that of $\vec{K}_-$ by $2\pi/3$.

Thus under this discrete rotation

$$b_k \to b_{k-2n_1} , \tag{62}$$
$$b_{k-n_1} \to b_{k-n_1-n_2} , \tag{63}$$
$$b_{k-n_2} \to b_{k-n_1} , \tag{64}$$
$$a_k \to a_{k-n_1} , \tag{65}$$
$$a_{k+n_1} \to a_{k-2n_1+n_2} , \tag{66}$$
$$a_{k+n_2} \to a_{k-2n_1} . \tag{67}$$

Also we observe that

$$\vec{K}'_+ \cdot \vec{\delta}_3 = \vec{K}_- \cdot \vec{\delta}_1 \tag{68}$$
$$\vec{K}'_+ \cdot \vec{\delta}_1 = \vec{K}_- \cdot \vec{\delta}_2 \tag{69}$$
$$\vec{K}'_+ \cdot \vec{\delta}_2 = \vec{K}_- \cdot \vec{\delta}_3 \tag{70}$$

Substituting Eqs. (62)-(70) into Eqs. (19)-(20) we obtain

$$i\hbar \dot{a}_{k-n_1} = -t_h(b_{k-2n_1} e^{i\vec{K}_- \cdot \vec{\delta}_1} + b_{k-n_1-n_2} e^{i\vec{K}_- \cdot \vec{\delta}_2}$$
$$+ b_{k-n_1} e^{i\vec{K}_- \cdot \vec{\delta}_3}) + U a^*_{k-n_1} a_{k-n_1} a_{k-n_1} , \tag{71}$$
$$i\hbar \dot{b}_{k-2n_1} = -t_h(a_{k-n_1} e^{-i\vec{K}_- \cdot \vec{\delta}_1} + a_{k-2n_1+n_2} e^{-i\vec{K}_- \cdot \vec{\delta}_2}$$

$$+ a_{k-2n_1} e^{-i\vec{K}_- \cdot \vec{\delta}_3}) + U b^*_{k-2n_1} b_{k-2n_1} b_{k-2n_1} . \tag{72}$$

Now we redefine the index k: in Eq. (71), $k \to k+n_1$, while in Eq. (72), $k \to k + 2n_1$. Then

$$i\hbar \dot{a}_k = -t_h(b_{k-n_1} e^{i\vec{K}_- \cdot \vec{\delta}_1} + b_{k-n_2} e^{i\vec{K}_- \cdot \vec{\delta}_2}$$
$$+ b_k e^{i\vec{K}_- \cdot \vec{\delta}_3}) + U a^*_k a_k a_k \tag{73}$$
$$i\hbar \dot{b}_k = -t_h(a_{k+n_2} e^{-i\vec{K}_- \cdot \vec{\delta}_1} + a_{k+n_2} e^{-i\vec{K}_- \cdot \vec{\delta}_2}$$
$$+ a_k e^{-i\vec{K}_- \cdot \vec{\delta}_3}) + U b^*_k b_k b_k \tag{74}$$

To summarize, a rotation by $\pi/3$ which takes $\vec{K}_+ \to \vec{K}'_+$ is identical to the unrotated theory but with $\vec{K}_+ \to \vec{K}_-$. Note that the redefinition of the index k is different for the two equations because old and new primitive cells are related by a rotation. Rotating by $\pm 2\pi/3$ exchanges A and B sites once more and returns Ψ to its original configuration. Since Ψ has four components, the effect of making one full rotation of 2π is that the components acquire a net phase so that $\Psi \to -\Psi$. This *Berry phase* [2] endows Ψ with the characteristic double-valuedness of a genuine 4-spinor.

However, we must be cautious when discussing chirality and helicity, since we treat (2+1) dimensions and can use only the first two Pauli matrices. As in the (3+1) theory, one can define a *pseudo-chirality operator* in (2+1) dimensions, γ^5, as the product of the other four γ matrices. In the Weyl, or Chiral representation we have

$$\gamma^5 \equiv i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{-1} \end{bmatrix} , \tag{75}$$

where again the boldface indicates a 2×2 submatrix. This is the natural representation for Ψ in that the NLDE of Eq. (56) maps into this representation in a natural way: states of well defined chirality correspond to the upper and lower 2-spinors. Thus the upper and lower spinors

$$\Psi_+ \equiv \begin{bmatrix} \psi_+ \\ \mathbf{0} \end{bmatrix} , \quad \Psi_- \equiv \begin{bmatrix} \mathbf{0} \\ \psi_- \end{bmatrix} \tag{76}$$

are eigenfunctions of γ^5,

$$\gamma^5 \Psi_+ = \Psi_+ , \quad \gamma^5 \Psi_- = -\Psi_- . \tag{77}$$

The question of Poincaré covariance of the nonlinear Dirac equation remains. First, we check coordinate translations. The wave function is required to transform as

$$\Psi'(r_\mu + d_\mu) = \Psi(r_\mu) . \tag{78}$$

Thus the nonlinear terms which have two factors of Ψ are invariant under translations. Under spatial rotations we observe that the interaction terms in the NLDE remain unchanged within the context of the full theory. We include both $\vec{K}_+$ and $\vec{K}_-$ points in the full theory. As for the case of boosts, we note that for the linear equation the components of the wave function transform in accordance with the transformation of the space-time coordinates in their

arguments. This is exactly canceled by the reciprocal transformations of the partial derivatives. This is not the case for the nonlinear terms. One has some matrix product with two factors of the wave function. Thus the *nonlinear* Dirac equation is not invariant under Lorentz boosts.

For any truly fundamental theory of nature it is demanded that the governing equations be invariant with respect to the Poincaré group. Sometimes this requirement is loosened as in the quantization of gauge theories, when fixing the gauge causes relativistic invariance to be non-manifest. Yet the theory itself certainly remains invariant. Other times the breaking of Poincaré covariance implies that we are dealing with an effective theory in which deeper physical processes are at work. Our case is an example of the latter. By including on-site interactions we break relativistic covariance of the linear equation of motion by introducing self-interactions which can be viewed as evidence of "deeper physics" in our two-dimensional universe.

3.3. *Hermiticity*

We require the Hamiltonian be Hermitian in order to guarantee that physically measurable quantities (observables) are real. Thus each term must be independently Hermitian. In particular, we must show that $N^\dagger = N$. The proof follows:

$$
\begin{aligned}
N^\dagger &= (U\Psi^\dagger A_+ \Psi A_+)^\dagger \\
&= U A_+^\dagger \Psi^\dagger A_+^\dagger (\Psi^\dagger)^\dagger \\
&= U A_+ \Psi^\dagger A_+ \Psi \\
&= U \Psi^\dagger A_+ \Psi A_+ \\
&= N
\end{aligned}
\tag{79}
$$

where the last two steps work because A_+ is real and symmetric. The nonlinear terms, and therefore the NLDE, are indeed Hermitian.

3.4. *Current Conservation*

Current conservation is expected in ordinary QED for a closed system, i.e., an isolated volume of space. However, most theories of quantum gravity do introduce Lorentz violating terms which bring in charge non-conservation effects. The conserved current for the linear Dirac equation is

$$
j^\mu = \bar{\Psi}\gamma^\mu\Psi
\tag{80}
$$

where $\bar{\Psi} \equiv \Psi^\dagger\gamma^0$.

We check that current is also conserved for the NLDE. The 3-divergence of the current in (2+1) spatial dimensions is

$$
\begin{aligned}
\partial_\mu j^\mu &= \partial_\mu(\bar{\Psi}\gamma^\mu\Psi) \\
&= \partial_\mu(\Psi^\dagger\gamma^0\gamma^\mu\Psi) \\
&= -(\partial_\mu\Psi^\dagger)\gamma^\mu\gamma^0\Psi + \Psi^\dagger\gamma^0\gamma^\mu\partial_\mu\Psi
\end{aligned}
\tag{81}
$$

where $\mu \in \{0,1,2\}$ and we have used the chain rule and the anti-commuting properties of the Dirac matrices. Taking the adjoint of the NLDE in four-component form, Eq. (56), yields

$$
(i\gamma^\mu\partial_\mu\Psi)^\dagger = (-\gamma^0 N\Psi)^\dagger ,
\tag{82}
$$

which implies

$$
-\partial_\mu\Psi^\dagger\gamma^\mu = -i\Psi^\dagger N^\dagger\gamma^0 .
\tag{83}
$$

Then the (2+1)-divergence of the current is

$$
\partial_\mu j^\mu = -i\Psi^\dagger N^\dagger\gamma^0\gamma^0\Psi + \Psi^\dagger\gamma^0(i\gamma^0 N\Psi)
\tag{84}
$$

$$
= -i\Psi^\dagger(N^\dagger - N)\Psi = 0
\tag{85}
$$

Thus current is conserved. This is in fact the statement that hermiticity implies current conservation.

3.5. *Chiral Current*

The conserved chiral current for the linear Dirac equation is

$$
j_5^\mu \equiv \bar{\Psi}\gamma^\mu\gamma_5\Psi .
\tag{86}
$$

Then the (2+1)-divergence of the chiral current is

$$
\partial_\mu j_5^\mu = \partial_\mu(\bar{\Psi}\gamma^\mu\gamma_5\Psi) .
\tag{87}
$$

Following a similar route as in Sec. 3.4, we find

$$
\partial_\mu j_5^\mu = -i\Psi^\dagger(N\gamma_5 + \gamma_5 N)\Psi
\tag{88}
$$

where we have used the hermiticity of N. This means that in order for chiral current to be conserved we must have

$$
\{N, \gamma_5\} = 0
\tag{89}
$$

where the curly braces signify the anti-commutator. The anti-commutator is

$$
\{N_+, \gamma_5\} = U\Psi^\dagger A_+ \Psi A_+\gamma_5 + \gamma_5 U\Psi^\dagger A_+ \Psi A_+ .
\tag{90}
$$

Writing this out in explicit matrix form we find that

$$
\{N_+, \gamma_5\} = +2N ,
\tag{91}
$$

$$
\{N_-, \gamma_5\} = -2N .
\tag{92}
$$

Therefore chiral current is not conserved. This fact is reminiscent of the anomalous non-conservation of chiral current in the case of some field theories upon quantization, which is mediated by instantons. It is interesting to consider that our nonlinear terms might be treated as quantum-induced nonlinearities.

3.6. *Universality*

Universality refers to the invariance of a theory under rescaling of the solution. For the case of the linear Dirac equation, $\Psi \to \lambda\Psi$ leaves the theory unchanged. Because

the nonlinear term in the NLDE contains one factor of Ψ and one factor of its adjoint, it scales as λ^2, thus breaking the invariance of the nonlinear theory. For a treatment of NLDEs which are universal in this sense, but not relevant to the present solid state system, see Ref. [15].

3.7. Discrete Symmetries

3.7.1. Parity

In the case of the standard massless Dirac equation, invariance under the parity operation $\hat{P}$ requires Ψ to transform as

$$\Psi \to \Psi' = \hat{P}\Psi = \gamma^0 \Psi \,, \tag{93}$$

where, in contrast to the Chiral representation presented in Eq. (54), we take γ^0 in this section to be in the Dirac representation,

$$\gamma^0 = \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{bmatrix} \tag{94}$$

We proceed to determine if the NLDE remains invariant under the transformation of Eq. (93). The parity operator acting on the honeycomb lattice inverts both coordinate axes, $\vec{\nabla} \to -\vec{\nabla}$, and thus exchanges A and B sites while also exchanging $\vec{K}$-point indices. The transformed *linear* equations ($U = 0$) are

$$i\gamma^{\mathbf{0}} \begin{bmatrix} \dot{\Psi}_- \\ \dot{\Psi}_+ \end{bmatrix} + ic_s \begin{bmatrix} -\vec{\sigma} \cdot \vec{\nabla} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \cdot \vec{\nabla} \end{bmatrix} \gamma^0 \begin{bmatrix} \Psi_- \\ \Psi_+ \end{bmatrix} = 0 \,. \tag{95}$$

Interchanging upper and lower spinors we obtain the equivalent form

$$i \begin{bmatrix} \dot{\Psi}_+ \\ \dot{\Psi}_- \end{bmatrix} + ic_s \begin{bmatrix} \vec{\sigma} \cdot \vec{\nabla} & \mathbf{0} \\ \mathbf{0} & -\vec{\sigma} \cdot \vec{\nabla} \end{bmatrix} \begin{bmatrix} \Psi_+ \\ \Psi_- \end{bmatrix} = 0 \,. \tag{96}$$

Thus the linear equations are invariant under parity.

Next we check the nonlinear term in the NLDE. With the transformed nonlinear term added to Eq. (95) we obtain

$$i \begin{bmatrix} \dot{\Psi}_- \\ -\dot{\Psi}_+ \end{bmatrix} + ic_s \begin{bmatrix} -\vec{\sigma} \cdot \vec{\nabla} & \mathbf{0} \\ \mathbf{0} & \vec{\sigma} \cdot \vec{\nabla} \end{bmatrix} \begin{bmatrix} \Psi_- \\ -\Psi_+ \end{bmatrix} - U \begin{bmatrix} N_- \\ -N_+ \end{bmatrix} = 0 \,, \tag{97}$$

which after similar steps as above gives us back the NLDE. Therefore parity is a symmetry of the nonlinear Dirac equation. We note that parity is conserved in standard QED and also in the theory of the strong interactions, but violated by weak interactions. Therefore, in this aspect, the NLDE is consistent with QED.

3.7.2. Charge Conjugation

In order to maintain symmetry under charge conjugation one requires the nonlinear term to transform as

$$(\gamma^0 N)' = \hat{C}\gamma^0 N^* \hat{C}^{-1} \tag{98}$$

where in the Chiral representation the charge conjugation operator is

$$\hat{C} = \gamma^1$$

and the wavefunction transforms in the standard way:

$$\Psi \to \gamma^1 \gamma^0 \Psi^{\dagger T} \,. \tag{99}$$

We determine if the NLDE maintains this symmetry. Let F be one of the nonlinear terms. Then

$$\begin{aligned} (\gamma^0 N)' &= \left(\gamma^1 \gamma^0 \Psi^{\dagger T}\right)^{\dagger} A_+ (\gamma^1 \gamma^0 \Psi^{\dagger T}) A_+ \\ &= \Psi^T \gamma^0 \gamma^1 A_+ \gamma^1 \gamma^0 \Psi^* A_+ \end{aligned} \tag{100}$$

Also, we have

$$\begin{aligned} \hat{C}\gamma^0 N^* \hat{C}^{-1} &= \gamma^1 \gamma^0 (\Psi^{\dagger} A_+ \Psi A_+)^* \gamma^1 \\ &= \gamma^1 \gamma^0 \Psi^T A_+ \Psi^* A_+ \gamma^1 \end{aligned} \tag{101}$$

Thus the NLDE breaks charge conjugation symmetry.

3.7.3. Time Reversal

The usual time reversal, $t \to -t$, requires that Ψ transform as

$$\Psi \to \Psi' = \hat{\Theta}\Psi = i\gamma^1 \gamma^3 \Psi \,, \tag{102}$$

where $\hat{\Theta}$ is the time-reversal operator and

$$i\gamma^1 \gamma^3 = \begin{bmatrix} -\sigma_y & \mathbf{0} \\ \mathbf{0} & -\sigma_y \end{bmatrix} \,, \tag{103}$$

In our theory the intrinsic effect of time reversal is to change the direction of momentum so that $\vec{K}$ points are switched without exchanging A and B indices. Combining these effects, we determine whether or not the linear part of the NLDE, i.e., the linear Dirac equation, remains invariant:

$$-i\frac{\partial}{\partial t} i\gamma^{\mathbf{1}}\gamma^{\mathbf{3}} \begin{bmatrix} \Psi_- \\ \Psi_+ \end{bmatrix} + ic_s \begin{bmatrix} \vec{\sigma} \cdot \vec{\nabla} & \mathbf{0} \\ \mathbf{0} & -\vec{\sigma} \cdot \vec{\nabla} \end{bmatrix} i\gamma^{\mathbf{1}}\gamma^{\mathbf{3}} \begin{bmatrix} \Psi_- \\ \Psi_+ \end{bmatrix} = 0 \tag{104}$$

Then

$$-i\frac{\partial}{\partial t} \begin{bmatrix} \Psi_+ \\ \Psi_- \end{bmatrix} + ic_s \begin{bmatrix} \vec{\sigma} \cdot \vec{\nabla} & \mathbf{0} \\ \mathbf{0} & -\vec{\sigma} \cdot \vec{\nabla} \end{bmatrix} \begin{bmatrix} \Psi_+ \\ \Psi_- \end{bmatrix} = 0 \tag{105}$$

The appearance of the negative sign in front of the time derivative indicates that these are the equations satisfied

by the negative energy solutions, or holes. Thus time reversal takes the equations describing electrons into holes and those for holes into electrons but keeps the overall theory invariant.

We proceed to consider the nonlinear term. With the transformed nonlinear term we obtain

$$
-i\frac{\partial}{\partial t}\begin{bmatrix}\Psi_+ \\ \Psi_-\end{bmatrix} + ic_s\begin{bmatrix}\vec{\sigma}\cdot\vec{\nabla} & \mathbf{0} \\ \mathbf{0} & -\vec{\sigma}\cdot\vec{\nabla}\end{bmatrix}\begin{bmatrix}\Psi_+ \\ \Psi_-\end{bmatrix}
$$
$$
-iU\begin{bmatrix}N_+ \\ N_-\end{bmatrix} = 0 \,. \tag{106}
$$

In Eq. (106) the interaction term has acquired a factor of i. Therefore the NLDE is not invariant under time reversal. Thus, as one also observes in the Standard Model of particle physics, CP and T are not conserved independently but CPT is conserved.

4. Discussion and Conclusions

The nonlinear Dirac equation we have presented introduces a completely new class of nonlinear phenomena in Bose-Einstein condensates. Although our work is related to graphene, in that the BEC is taken to be trapped on a honeycomb lattice, we have switched bosons for fermions. The form of the nonlinearity is then a natural physical result of binary interactions between bosons. In fact, it is a spinor generalization of the kind of nonlinearity one finds in the nonlinear Schrodinger equation, i.e., proportional to the local condensate density. The same equation will occur for light subject to a Kerr nonlinearity when propagating through a photonic crystal with a honeycomb lattice structure [33].

We showed that the NLDE breaks Poincaré covariance, and therefore the principle of relativity. This suggests that small nonlinearities of this form could be looked for by such symmetry breaking in a variety of systems where Dirac or Dirac-like equations apply. For instance, even for fermions there is a small mean field effect. We could just as easily have considered ultra-cold fermions on an optical lattice. This would appear at first sight to be the exact analog of graphene; however, our work shows that Poincaré covariance will be broken by mean field effects, even if on a small level. Indeed, for graphene one should expect similar effects due to Coulomb interactions. The latter nonlinearity can be expected to be non-local due to the power law behavior of $1/r$ for the Coulomb potential, just as dipole-dipole interactions between ultra-cold atoms even without an optical lattice lead to a non-local nonlinear Schrodinger equation in a 3D continuum (in 2D dipole-dipole interactions, which have a power law of $1/r^3$, are local). Thus, in graphene, we make the conjecture that there is a non-local nonlinear correction term to the massless Dirac equation which breaks Poincaré covariance. Similar corrections can be looked for in quantum electrodynamics under the proper circumstances [21].

All generalizations of the scalar nonlinear Schrodinger equation relevant to BECs are candidates for generalized nonlinear Dirac equations. For example, pseudo-spin structure leads to a vector NLSE. One can therefore anticipate a vector NLDE as well.

We have thoroughly explored both continuous and discrete symmetries of the NLDE. In particular, we showed that pseudo-chiral current is not conserved, the NLDE is not covariant under Lorentz boosts, and it breaks charge conjugation as well as time reversal symmetry. On the other hand, the NLDE is hermitian, local, conserves current, and is symmetric under parity. The NLDE maintains CPT symmetry, just as one finds in the Standard Model.

In a future work we will treat soliton and vortex solutions of the NLDE, as a first step towards a complete classification of nonlinear relativistic phenomena in BECs.

References

[1] K. S. Novoselov, A. K. Geim, S. V. Morozov, D. Jiang, M. I. Katsnelson, I. V. Grigorieva, S. V. Dubonos, A. A. Firsov, Nature 438 (2005) 197.
[2] Y. B. Zhang, Y. W. Tan, H. L. Stormer, P. Kim, Nature 438 (2005) 201–204.
[3] A. K. Geim, K. S. Novoselov, Nature Materials 6 (2007) 183.
[4] C. Wu, D. Bergman, L. Balents, S. D. Sarma, Phys. Rev. Lett. 99 (2007) 070401.
[5] Z. Hadzibabic, P. Kruger, M. Cheneau, B. Battelier, J. B. Dalibard, Nature 441 (2006) 1118.
[6] S.-L. Zhu, B. Wang, L.-M. Duan, Phys. Rev. Lett. 98 (2007) 260402.
[7] F. Dalfovo, S. Giorgini, L. P. Pitaevskii, S. Stringari, Rev. Mod. Phys. 71 (1999) 463.
[8] A. J. Leggett, Rev. Mod. Phys. 73 (2001) 307.
[9] P. G. Kevrekidis, D. J. Frantzeskakis, R. Carretero-González (Eds.), Emergent Nonlinear Phenomena in Bose-Einstein Condensates, Springer-Verlag, Berlin, 2008.
[10] R. V. Mishmash, L. D. Carr, Mathematics and Computers in Simulation, under review (2008).
[11] R. V. Mishmash, L. D. Carr, e-print arXiv:0710.0045.
[12] W. Fushchich, R. Zhdanov, Physics Reports 172 (1989) 123.
[13] B. I. F.M. Toyama, Y. Hosona, Y. Nogami, J. Phys. A: Math. Gen. 27 (1994) 3139.
[14] L. Z. Hu, L. Y. Hu, Int. J. Theor. Phys. 37 (1998) 2115.
[15] W.-K. Ng, R. R. Parwani, (2007) e-print arXiv.org:0707.
[16] T. Cazenave, L. Vazquez, Commun. Math. Phys. 105 (1986) 35.
[17] M. J. Esteban, E. Seré, Commun. Math. Phys. 171 (1995) 323.
[18] W. Fushchych, R. Zhdanov, Symmetries And Exact Solutions Of Nonlinear Dirac Equations, Mathematical Ukraina Publisher, Ukraine, 1997.
[19] M. J. Esteban, E. Seré, Discrete and Continuous Dynamical Systems 8 (2002) 381.
[20] A. Maccari, Phys. Lett. A 336 (2005) 117.
[21] A. Ferrando, H. Michinel, M. Seco, D. Tommasini, Phys. Rev. Lett. 99 (2007) 150404.

[22] M. Katsnelson, K. S. Novoselov, e-print arXiv:cond-mat/0703374.

[23] A. M. Rey, Ultracold bosonic atoms in optical lattices, Ph.D. thesis, University of Maryland (2004).

[24] P. B. Blakie, C. W. Clark, J. Phys. B: At. Mol. Opt. Phys. 37 (2004) 1391.

[25] L. P. Pitaevskii, Sov. Phys. JETP 13 (1961) 451–454; E. P. Gross, Nuovo Cimento 20 (1961) 454.

[26] Y. Castin, in: R. Kaiser, C. Westbrook, F. David (Eds.), Coherent atomic matter waves, EDP Sciences and Springer-Verlag, Les Ulis, France and Berlin, Germany, 2001, pp. 1–136.

[27] P. R. Wallace, Phys. Rev. 71 (1947) 622.

[28] E. Taylor, E. Zaremba, Phys. Rev. A 68 (2003) 053611.

[29] B. Wu, J. Shi, e-print cond-mat/0607098.

[30] V. Gusynin, S. Sharapov, J. Carbotte, e-print arXiv:0706.3016v1.

[31] J. D. Bjorken, S. D. Drell, Relativistic Quantum Mechanics, McGraw-Hill, New York, NY, 1964.

[32] J. Hladik, Spinors in Physics, Springer-Verlag, 1999.

[33] N. M. C. Aközbek, S. John, Phys. Rev. E 57 (1998) 2287–2319.